Recent Advances in Chemoenzymatic Bioconjugation Methods

Jesse M. McFarland and David Rabuka

Catalent Biologics–West, Catalent Pharma Solutions, Emeryville, CA, USA.

ABSTRACT: Genetic fusions of either full enzymes or peptide tags with a protein of interest have enabled the synthesis of protein conjugates with precise control over the site of attachment and number of payloads incorporated. Engineering of protein glycans, depending on the application, can lead to similar control for nonengineered glycoproteins. Recent advances in the field of chemoenzymatic modifications of proteins for site-specific protein conjugation will be reviewed. These techniques have been used in an array of fields for the execution of innovative and valuable experiments. Specific industrial applications of these technologies will be highlighted.

KEYWORDS: bioconjugation, antibody drug conjugate, protein tag, transferase, ligase, glycoconjugate, chemoenzymatic

CITATION: McFarland and Rabuka. Recent Advances in Chemoenzymatic Bioconjugation Methods. *Organic Chemistry Insights* 2015:5 7–14 doi:10.4137/OCI.S17957.

TYPE: Review

RECEIVED: May 20, 2015. **RESUBMITTED:** June 21, 2015. **ACCEPTED FOR PUBLICATION:** June 29, 2015.

ACADEMIC EDITOR: Yitzhak Tor, Editor in Chief

PEER REVIEW: Four peer reviewers contributed to the peer review report. Reviewers' reports totaled 374 words, excluding any confidential comments to the academic editor.

FUNDING: Authors disclose no funding sources.

COMPETING INTERESTS: The authors are employees of Catalent Pharma Solutions.

CORRESPONDENCE: david.rabuka@catalent.com

Paper subject to independent expert blind peer review. All editorial decisions made by independent academic editor. Upon submission manuscript was subject to anti-plagiarism scanning. Prior to publication all authors have given signed confirmation of agreement to article publication and compliance with all applicable ethical and legal requirements, including the accuracy of author and contributor information, disclosure of competing interests and funding sources, compliance with ethical requirements relating to human and animal study participants, and compliance with any copyright requirements of third parties. This journal is a member of the Committee on Publication Ethics (COPE). Provenance: the authors were invited to submit this paper.

Published by Libertas Academica. Learn more about this journal.

Introduction

As the field of Chemical Biology has evolved in the last two decades, the use of enzymes to make specific covalent bonds leading to biomolecules with defined chemical composition has rapidly expanded. The drive for the development of new technologies is motivated by a variety of needs, both industrial and academic. Growing emphasis on homogeneous therapeutic protein conjugates is in recognition of the benefits conferred over nonspecific conjugates and a major force in the development of in vitro technologies. Cell applications have motivated the development of technologies that enable experiments, particularly microscopy-based experiments, which cannot be accomplished with green fluorescent protein (GFP)–based reagents.

Strategies to prepare site-specific protein conjugates by chemoenzymatic techniques generally fall into three categories. The first incorporates an enzyme as a fusion with the protein of interest that reacts specifically with a synthetic reagent bearing a payload (often a fluorophore, affinity handle, or cytotoxin). The second uses a small peptide tag incorporated into the protein that is recognized by an enzyme, which transfers a payload to the target. The third strategy takes advantage of the rich field of carbohydrate chemistry to remodel the glycan on a protein and thereby introduce a payload (Fig. 1A–C). In each strategy, a second synthetic step is required if the chemoenzymatic step introduces a reactive handle, such as an azide or alkyne, that is primed for further elaboration. While a second step may appear disadvantageous, often a significant excess of the primary substrate is required relative to the protein target, which can be prohibitive if the ultimate payload is exceptionally expensive and/or difficult to prepare (Fig. 1D).

This review will focus on the use of enzymes for preparing small molecule–protein conjugates and the advances made in the last 5 years, both in new techniques and their applications. Several reviews cover earlier work.[1–5] Where possible, we will highlight current industrial applications.

Enzyme–protein Fusions

Recombinant fusions of peptide tags (eg, FLAG, HA, or myc) or the GFP and related protein fluorophores have been widely used due to their ease of preparation. Indeed, the Nobel Prize Committee recognized the broad importance of GFP in 2008. However, GFP has several drawbacks, including its size (27 kD), biophysical properties (dimers often observed by SDS-PAGE), and rapid photobleaching when used in advanced microscopy techniques.

Inspiration by natural enzymes that recognize and bind unusual chemical motifs has led to the development of a collection of technologies including SNAP- and CLIP-tags, Halo-Tag, dihydrofolate reductase–trimethoprim (DHFR-TMP) tag, fluorogen-activating proteins (FAPs), and cutinase[6,7] (not discussed here), which overcome some drawbacks of using GFP.

Johnsson and coworkers developed SNAP-tag, a small enzyme (20 kD) that recognizes O^6-benzylguanine–containing substrates and transfers the payload to an active-site cysteine.[8,9] The same group later introduced the CLIP-tag (21 kD), which uses a complementary O^2-benzyl cytosine substrate.[10] Trono and coworkers engineered SNAP-tag fusions

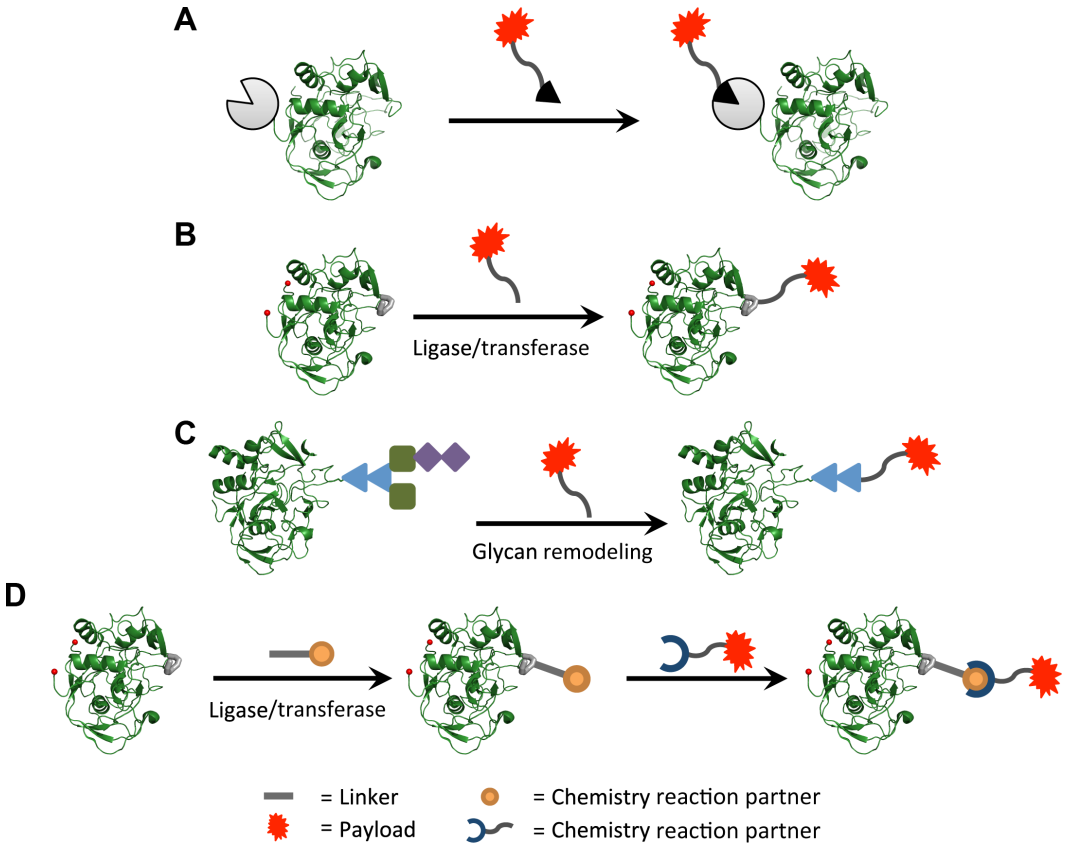

Figure 1. General strategies for chemoenzymatic modification of proteins. (**A**) A protein of interest is genetically fused to an enzyme that recognizes a specific substrate and covalently modifies the enzyme. (**B**) A tag sequence (in gray) is introduced into the protein, typically in a flexible loop, which is recognized by an enzyme that links a payload to the tag. Some enzymes also specifically recognize the N- or C-terminus (red spheres) and link the payload at that position. (**C**) One or multistep glycan remodeling gives a conjugate with the payload linked to the modified glycan. (**D**) An example of two-step protein labeling in which the payload of interest is attached in a second chemical step after introduction of a chemical handle.

with both CD4 and KAP1 and showed that the SNAP-tag rapidly reacted with its benzyl guanine partner in vivo and that pulse chase experiments could determine protein half-life in vivo.[11] Subsequently, Heppenstall and coworkers engineered a Cre-dependent reporter mouse in which the SNAP-tag was fused to a membrane-targeting CAAX consensus sequence for Cre-dependent expression. The fusion was found to react with a benzyl guanine partner both in vivo and ex vivo.[12]

Another popular tag is the HaloTag (34 kD), which recognizes an alkyl chloride substrate and transfers the alkyl group to the active-site aspartate.[13] Recently, Crews and coworkers used the HaloTag in combination with an unusual substrate that destabilizes the tertiary structure of the tag. When fused to an ER-targeting signal, substrate-induced unfolding transiently activated the unfolded protein response (UPR). This strategy allowed for the identification of estrogen receptor–dependent signaling in the UPR that is typically masked by pro-apoptotic signaling induced by commonly used reagents like thapsigargin.[14] Piehler and coworkers used a HaloTag fusion for the micro-patterning of a membrane-presented type I interferon receptor. Treatment of the cells with interferon enabled the observation of interferon receptor dynamics, recruitment of JAK/STAT proteins, and subsequent downstream signaling and negative feedback of the pathway by USP18.[15]

A number of examples have appeared in which a dual-tag system is employed. Maly and coworkers designed a trifunctional linker for activity-based protein-profiling experiments. The linker comprised both SNAP and Halo tag domains as well as a protease cleavage site; target-bound bait (a type II kinase inhibitor) reacted with the HaloTag and was then pulled down with a benzyl guanine–functionalized resin via the SNAP-tag. A panel of kinases, including several unexpected members, bound the bait molecule, demonstrating the challenges of correlating in vitro activity with in vivo activity, and especially in identifying off-target interactions.[16]

In certain applications, a desirable property is the turn on of fluorescence only after a nonfluorescent ligand binds its target protein. Cornish and coworkers have developed a system based on the specific binding of TMP to DHFR (18 kD). A TMP-fluorophore-quencher molecule rapidly reacts with an engineered DHFR in live cells, displacing the quencher and turning on fluorescence. The method was demonstrated with several cellular substrates and shown to be effective for live cell imaging in a proof-of-concept experiment.[17] Waggoner

and coworkers have developed an alternative system termed fluorogen-activating protein (FAP; 14–30 kD, depending on specific substrate pair) in which scFvs has been evolved to bind Malachite Green and other dyes.[18] Bruchez and coworkers have similarly demonstrated that FAP can be used for no-wash live cell imaging.[19] Lidke and coworkers fused an FAP domain to the gamma chain of FcεRI to measure receptor mobilization and internalization. Single-particle tracking of the labeled gamma chain enabled the observation that receptor dissociation did not occur prior to or during receptor endocytosis after activation with a low or high cytokinergic IgE variant and that IgE binding to the receptor did not alter FcεRI mobility in the absence of a cross-linking antigen.[20]

While these protein-based tags are not much different in size from GFP (25 kD), they offer other advantages including choices of different dyes or payloads (commercially available from a variety of companies)[21] to pair with each tag and greater control over how and when the fluorophore is activated.

Peptide Tags for Protein Labeling

A variety of transferase or ligase enzymes have been identified in recent years and repurposed for site-specific protein modification. Typically, a small peptide sequence incorporated into the target of interest is recognized by the enzyme and then transfers a payload from an analog of its natural substrate onto the tag. Examples include Sortase A (SrtA), protein farnesyl transferase (PFTase), phosphopantetheinyl transferase (PPTase), bacterial transglutaminase (BTGase), biotin ligase, lipoic acid ligase (LAL), *N*-myristoyl transferase (NMTase), glutathione *S*-transferase (GST), SpyTag, and several engineered protease transferases. A primary benefit of this approach is the small size of the peptide tag that must be incorporated, which ranges from 3 to 15 residues. Some enzymes only recognize the tag peptide at a specific position in the primary sequence of the protein (often the N- or C-terminus), while others are not inherently limited by tag position.

Several ligases have been developed based on lipid or coenzyme transferases. The first such example was the biotin ligase, BirA, developed by Ting and coworkers.[22] The enzyme was found to transfer a biotin analog substrate containing a ketone onto the acceptor peptide sequence (15 residues). The method is limited by BirA substrate requirements,[23] which lead the same group to identify LAL as an alternative (13-residue tag sequence).[24] Subsequent work identified an LAL mutant that tolerates azide- or fluorophore-containing substrates.[25,26]

Another coenzyme transferase utilized for protein conjugation is the PPTase Sfp, initially reported by Walsh and coworkers.[27] Sfp binds synthetic analogs of acetyl CoA, catalyzing the transfer of the phosphopantetheine moiety to an 11-residue tag. In a recent report, Handel and coworkers applied the method to the labeling of a selection of chemokines with fluorophores for microscopy.[28] Novartis has applied the PPTase technology to the preparation of site-specific antibody–drug conjugates (ADCs).[29] Introduction of the tag sequence into the constant region of the heavy chain of the α-Her2 antibody trastuzumab facilitated the ligation of a cytotoxic payload (Fig. 2A). Depending on the position of the tag in the overall sequence, the structure of the antibody was not altered, as measured by melting temperature. Also, the resulting ADCs were potent in an in vivo xenograft study.

Two other lipid transferase enzymes have been used for protein conjugation: PFTase and NMTase. Recently Distefano and coworkers employed PFTase for the affinity purification of proteins. An aldehyde-containing farnesyl pyrophosphate substrate is transferred to the CAAX-box recognition sequence in cell lysates. Pull down of the labeled protein is then accomplished with hydrazide-functionalized resin. The captured protein can then be released from the resin by trapping with an amino-oxy–functionalized payload.[30,31] Work to improve the kinetics of PFTase resulted in a mutant with up to 300-fold increased rate for transfer of a coumarin-containing FPP analog.[32] Legochem Biosciences is working to commercialize the PFTase technology for the production of ADCs (Fig. 2B).[33] While PFTase only recognizes the CAAX-box at the C-terminus of proteins, NMTase acylation is limited to the N-terminal amine.[34,35] A report by Tirrell and coworkers shows the ease of microprinting of an N-terminally modified protein directly from *Escherichia coli* lysate.[36]

Transglutaminase enzymes catalyze the formation of an isopeptide bond between glutamine side chain amides and ε-amine of lysine residues. Many groups have utilized BTGase to recognize modified alkyl amines and transfer them to glutamine residues on the protein of interest. BTGase has two key characteristics that make it useful for bioconjugation: first, the enzyme recognizes a wide variety of alkyl amine substrates; second, BTGase has strict, though not yet defined, requirements for the recognition of the partner glutamine residue. If the alkyl amine is suitably functionalized, either the desired payload can be introduced directly or through a second bioorthogonal chemical step. BTGases have found a unique application in the development of ADCs because only a single Gln (Gln295) on the antibody heavy chain is recognized as a substrate after deglycosylation of the adjacent Asn297 (Fig. 2C).[37] This enables the production of highly homogenous ADCs.[38] Innate Pharma has introduced the mutation Asn297Gln, which eliminates the glycosylation site and leads to the production of ADCs with two drugs per heavy chain, as BTGase also recognizes Gln297 as a substrate.[39] Strop and coworkers at Rinat/Pfizer (South San Francisco, USA) have developed a four-residue glutamine tag sequence (LLQG) that can be introduced in any flexible and surface-accessible position in the primary sequence (Fig. 2A).[40] ADCs produced with the tag at different locations in the heavy or light chains display remarkably different biophysical, in vitro, and in vivo properties despite having the same overall chemical composition, highlighting the importance of choosing the payload conjugation site carefully.[41] Rinat/Pfizer has successfully moved an ADC prepared via this method into Phase I clinical trials.[42]

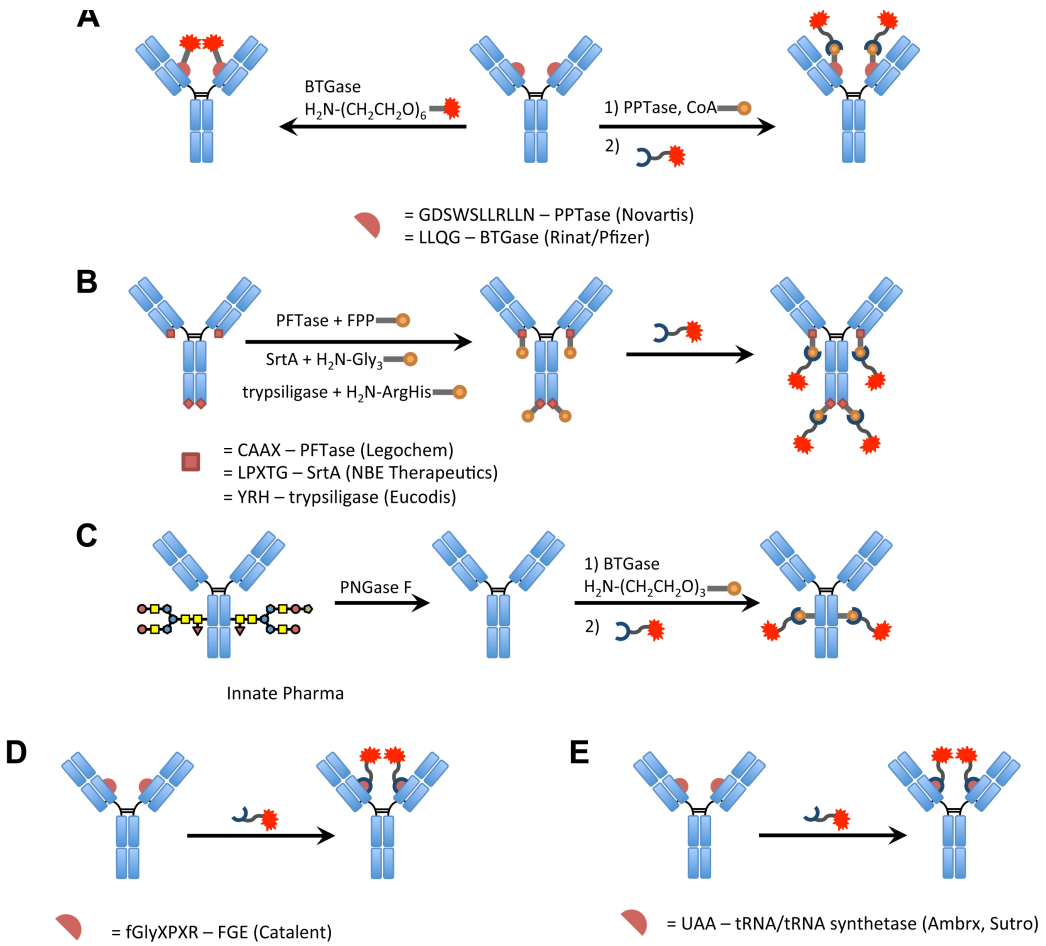

Figure 2. Industrial examples of chemoenzymatic strategies using small peptide tags for the production of ADCs. Companies working on each technology identified in parentheses. (Refer to Figure 1 legend for symbol descriptions.) (**A**) Different peptide tags recognized by PPTase or BTGase can be incorporated into any position in the antibody sequence. (**B**) Different tags that can only be incorporated at the protein C-terminus. (**C**) Removal of the glycan allows for Gln295 to be targeted by BTGase. (**D**) FGE converts cysteine to fGly, which can then be targeted with chemistry. (**E**) Unnatural amino acids can be incorporated during protein expression and then targeted with subsequent chemistry.

The ligation of a peptide-based substrate to form a native peptide bond can be accomplished by the use of SrtA. Plough and coworkers first disclosed the use of SrtA for protein bioconjugation, where SrtA recognizes a five-residue sequence (LPXTG) and performs a transamidation with a glycine nucleophile.[43,44] Schmohl and Schwarzer have reviewed SrtA in depth.[45] Pentelute and coworkers developed a method by which SrtA catalyzes the introduction of a thioester to a target protein. The product is then primed for participation in a native chemical ligation reaction, enabling the semisynthesis of, for example, GFP and anthrax toxin.[46] Donnelly and coworkers have used SrtA to introduce a Cu-chelating sarcophagine into an scFv targeting the active conformation of GPIIb/IIIa, a marker of platelet activation.[47] Positron emission tomography (PET) imaging of a murine model of carotid artery thrombosis showed specific uptake of the radiotracer at the site of injury.[48] One particular challenge of working with SrtA is that the peptide product of the ligation reaction also contains the recognition sequence for the enzyme, leading to product consumption over time. Several strategies have emerged to improve the usability of SrtA. Pentelute and coworkers designed a flow-based process that allows the use of lower concentrations of nucleophile without degradation of product yield.[49] Liu and coworkers chose to evolve SrtA using a yeast display technology, which lead to a mutant with 140-fold greater activity over wild type.[50] A subsequent evolution project resulted in two new SrtA mutants with altered sequence specificity (LPXSG and LAXTG).[51] Rather than release a glycine that can participate in the reverse reaction, Turnbull and coworkers have shown that SrtA can accept a depsipeptide substrate, which releases an alcohol that cannot participate in the reaction.[52,53] Although this strategy is limited to modifying the N-terminus, the synthetic depsipeptide allows the reaction to be completed within hours without a large excess of the nucleophile partner. Finally, Chilkolti and coworkers have identified isopeptide enzymatic activity by SrtA. While other sortase family members are known to catalyze isopeptide bond formation in their native context, this is the first report using SrtA with engineered substrates.[54] This allows for the conjugation of payloads to a tag inserted at any position within the primary sequence of a protein. Moreover, the product is not

a substrate for SrtA so there is no competition with the reverse reaction. On the industrial side, NBE Therapeutics (Basel, Switzerland) has implemented sortagging for the production of site-specific ADCs (Fig. 2B).[55]

Proteases cleave a protein backbone at sequence-specific or residue-specific positions. Engineering of a protease, its substrate, or both can promote the formation, as opposed to cleavage, of a peptide bond. Wild-type proteases used in combination with amino acid esters have successfully catalyzed the synthesis of polymers.[56,57] Bordusa and coworkers have designed a mutant of trypsin, termed trypsiligase, which recognizes the three-residue sequence YRH and cleaves after the tyrosine. A payload with an Arg-His ligation element is then acylated by the enzyme-acyl intermediate. Ligation of synthetic tags has been performed at both the N-terminus[58] and C-terminus[59] of proteins. While the formation of the product is fast and occurs under mild conditions at neutral pH, care must be taken, as the desired conjugate will also be degraded by trypsiligase. Eucodis has renamed the enzyme CTAT and has generated anti-Her2 ADCs with the payload at the antibody C-terminus (Fig. 2B).[60]

Howarth and coworkers have developed a rationally designed, split protein based on the fibronectin-binding protein FbaB, dubbed SpyTag. In one design, a small protein (15 kD) recognizes a 13-residue sequence and forms an isopeptide bond between a specific lysine–aspartate pair.[61] Tirrell, Arnold, and coworkers have used this technology to create a biopolymer incorporating a variety of bioactive proteins, including the leukemia inhibitory factor (LIF). The LIF-biopolymer successfully encapsulated murine embryonic stem cells and supported robust colony formation and maintenance of pluripotency.[62] Joshi and coworkers incorporated the SpyTag into their biofilm array technology and successfully incorporated recombinant α-amylase after biofilm formation.[63] The second-generation design permits for the ligation of two complementary peptide tags (11 and 13 residues, respectively) by an engineered SpyLigase enzyme.[64] The full potential of this technology remains to be explored.

Pentelute and coworkers have harnessed the promiscuity of GST to develop a method for rapidly labeling the glutathione thiol with perfluoroarenes in the presence of additional cysteine residues. In one example, GST catalyzed the macrocyclization of a 40-residue peptide in 70% crude yield in 2 hours.[47] Due to the requirement for both an N-terminal γ-Glu–Cys–Gly sequence and perfluoraryl reaction partner, this technology is currently limited to peptide-based couplings.

In contrast to the ligases, in which a payload bearing substrate is ligated to a recombinant tag sequence, formylglycine-generating enzyme (FGE) recognizes a 5-residue tag sequence (CXPXR) and then oxidizes the cysteine within the sequence to formylglycine (fGly).[65] A second chemical step is then required to introduce a payload of interest. Recently, Bertozzi and coworkers replaced the glycosylation site in an Fc domain with the FGE recognition sequence. Conjugation with an aminooxy GlcNAc enabled subsequent elaboration of the glycan with oxazoline glycans catalyzed by the EndoS mutant D233Q, providing a fully synthetic glycan.[66] Canonically, an oxime formed with the aldehyde moiety of fGly is an efficient bioconjugation strategy, but our laboratory has found this particular oxime to be hydrolytically unstable. To circumvent this problem, Agarwal, Kudirka and coworkers have developed new chemistries that form stable bonds with fGly, based on the Pictet–Spengler reaction[67] and the Knoevenagel reaction.[68] The combination of fGly and aldehyde-reactive chemistries are being developed by Catalent Biologics under the SMARTag™ brand (Fig. 2D). While several crystal structures of FGE have been published, Holder and coworkers have now shown that FGE binds a Cu^{2+} ion, which is required for catalytic activity and that the specific activity of recombinant FGE can be greatly enhanced by the addition of $Cu(II)SO_4$.[69]

Similar to FGE, mushroom-tyrosinase was discovered to selectively oxidize tyrosine residues in unstructured regions of proteins, such as in the HA tag.[70] The ortho-quinone intermediate can induce backbone cleavage or in the presence of a nucleophile, such as Besthorn's reagent (a benzothiazolinone hydrazone), can be trapped to yield a fluorescent adduct.

In the case of unnatural amino acid incorporation, an evolved tRNA synthetase charges the tRNA with the unnatural amino acid. The desired payload is typically introduced in a second step after protein expression and purification. The field has been recently reviewed in depth.[3,71] Two companies are working to commercialize the technology for the production of ADCs: Ambrx has incorporated a *p*-acetylphenylalanine-tRNA–tRNA synthetase pair into a CHO cell line for therapeutic protein expression and have successfully produced site-specifically conjugated ADCs targeting Her2[72] and CXCR4.[73] Sato and coworkers at Sutro Biopharma have developed a cell-free expression system to incorporate *p*-azidomethyl-phenylalanine. Subsequent conjugation with an alkyne-functionalized payload gives a homogeneous ADC (Fig. 2E).[74]

Glycan Remodeling

The potential of glycan remodeling has only been partially realized, as the high complexity of glycans and large families of glycosyl transferases have been daunting for development. For a glimpse of carbohydrate remodeling enzymes, visit CAZY on the web (http://www.cazy.org/). Recent reviews have highlighted advancements in the study of glycosyl transferases,[75] glycotherapies,[76] and methods for preparing glycoproteins for performing structure–function studies.[77–79]

Wu and coworkers developed a technology for the evaluation of LacNAc levels in histology sections. Taking advantage of the promiscuity of α1,3-fucosyltransferase in binding modified monosaccharides, the researchers were able to transfer an alkyne-labeled fucose to LacNAc-decorated glycans within formalin-fixed paraffin-embedded tissue samples, which were subsequently detected with a fluorescent click partner.[80]

Lewis and coworkers employed a similar strategy for the preparation of multimodal imaging conjugates. First, terminal galactose residues were removed from the glycans of an anti-A33 antibody by β1,4-galactosidase, followed by capping with a mutant galactotransferase and azide-containing GalNAc analog (aka GalNAz). Click conjugation with an Alexa Fluor 680 and a metal chelator lead to an immunoconjugate that enabled PET and fluorescent imaging of A33-positive xenograft tumors in mice.[81] A significant benefit of this strategy is that it does not require the engineering of a tag sequence into the antibody. However, it remains to be seen whether dramatic reengineering of ADC glycans will have unforeseen limitations in therapeutic applications. Wang and coworkers demonstrated the importance of controlling the glycoform of the Fc region by the chemoenzymatic synthesis of homogeneous glycans on Fc domains[82] and intact antibodies.[83] Surface plasmon resonance studies of different Fc–glycan conjugates demonstrated enhanced FcγRIIIa binding for bisected GlcNAc glycans.

While the potential for site-specifically modifying a protein without encoding a protein tag or domain is powerful, the methods are limited to proteins that are expressed in eurkaryotic cells, as bacterial expression systems lack the appropriate glycosylation machinery. Wang and coworkers have reported an initial strategy for the recombinant expression of eukaryotic *N*-glycoproteins in *E. coli*. The researchers transferred a selection of genes from the *Campylobacter jejuni pg1* locus into *E. coli* and took advantage of endogenous WecA, which catalyzes a key GlcNAc transfer to lipid P during lipid-linked oligosaccharide preparation prior to cotranslational transfer to the nascent protein. This combination produced a eurkaryotic Asn-linked GlcNAc glycan that could be elaborated in vitro.[84] DeLisa and coworkers revised this strategy to enable the production of trimannosyl chitobiose glycans in *E. coli*. By importing four yeast glycosyltransferases and one bacterial transferase from *C. jejuni*, the researchers were able to express a selection of proteins with the core eukaryotic N-linked glycan.[85] This technology is now being commercialized by Glycobia.[86]

A number of companies have attempted to harness glycoengineering for the production of therapeutic protein conjugates. Both Synaffix and Glykos use galactosyl transferases to introduce modified sugars onto deglycosylated antibodies for the subsequent attachment of a cytotoxic payload and production of ADCs (Fig. 3A).[87,88] Novo Nordisk has applied sialyl transferases to the production of protein–polymer conjugates for extending the half-life of therapeutic proteins such as factor VIIa (Fig. 3B).[89] Oxyrane, in collaboration with Callewaert and coworkers, have engineered a yeast expression system with a mannosidase (CcGH92_5) from *Arthrobacter luteus* that gives increased expression of mannose-6-phosphate–capped glycoproteins. Expression of α-glucosidase in this system, which is used for the treatment of Pompe disease, leads to more efficient targeting and uptake by the lysosome as compared to currently marketed therapeutic molecules (Fig. 3C).[90]

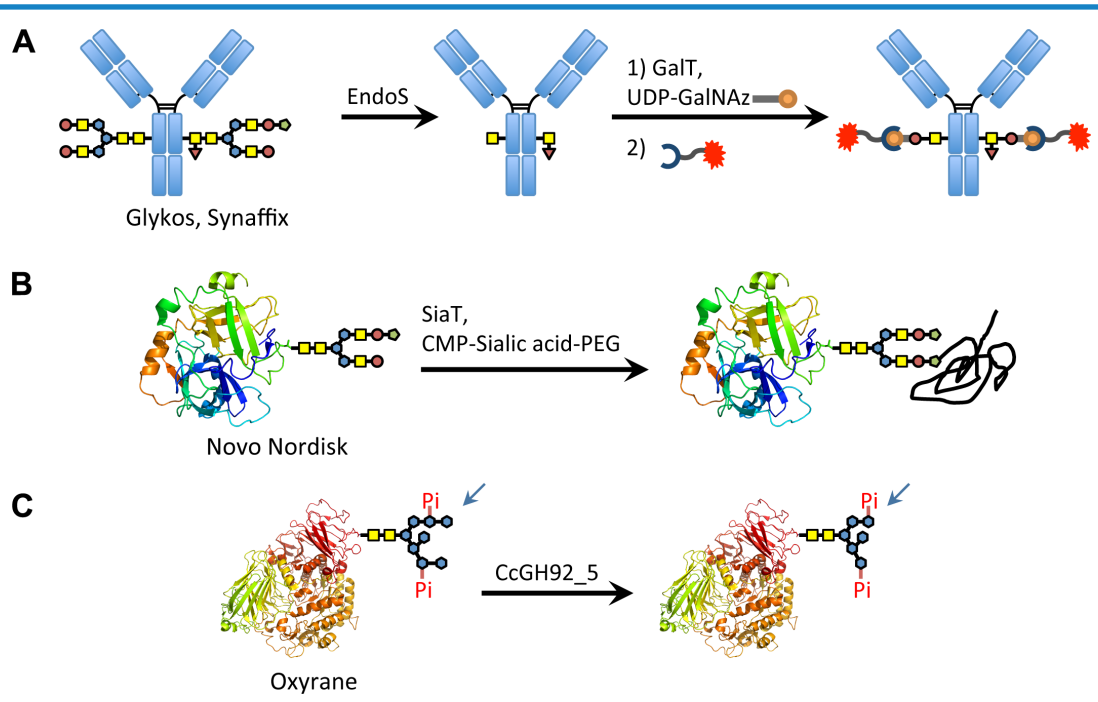

Figure 3. Examples of chemoenzymatic strategies targeting glycan remodeling for the production of therapeutic protein conjugates being commercialized. (Refer to Figure 1 legend for symbol descriptions.) (**A**) Glycan remodeling for ADC production. (**B**) Glycan remodeling for protein–polymer conjugation. (**C**) Glycan engineering followed by in vitro trimming yields a human glycoprotein that is efficiently targeted to the lysosome.

Conclusion

As the field of bioconjugation, and especially site-specific bioconjugation, has matured over the last decade, the number of techniques available to researchers has proliferated. Currently, the vast majority of ADCs in clinical trials are heterogenous conjugates, but the first site-specific, chemoenzymatic conjugate has now advanced to Phase I trials and we expect many more to enter the clinic in the near future. One significant challenge for many chemoenzymatic technologies will be the development of manufacturing processes that are scalable and cost-effective, although this appears to be a solvable problem. Looking forward, the biggest question will not be whether new chemoenzymatic techniques will be discovered, but how will researchers apply these tools to answer increasingly complex questions in fundamental biology and/or discover and develop protein conjugates for human benefit.

Author Contributions

Conceived and designed the experiments: DR, JMM. Analyzed the data: DR, JMM. Wrote the first draft of the manuscript: JMM. Contributed to the writing of the manuscript: DR. Agree with manuscript results and conclusions: DR, JMM. Jointly developed the structure and arguments for the paper: DR, JMM. Made critical revisions and approved final version: DR. Both authors reviewed and approved of the final manuscript.

REFERENCES

1. Rabuka D. Chemoenzymatic methods for site-specific protein modification. *Curr Opin Chem Biol*. 2010;14(6):790–796.
2. Rashidian M, Dozier JK, Distefano MD. Enzymatic labeling of proteins: techniques and approaches. *Bioconjug Chem*. 2013;24(8):1277–1294.
3. Lang K, Chin JW. Cellular incorporation of unnatural amino acids and bioorthogonal labeling of proteins. *Chem Rev*. 2014;114(9):4764–4806.
4. Yan Q, Bruchez MP. Advances in chemical labeling of proteins in living cells. *Cell Tissue Res*. 2015;360:179–194.
5. Agarwal P, Bertozzi CR. Site-specific antibody-drug conjugates: the nexus of bioorthogonal chemistry, protein engineering, and drug development. *Bioconjug Chem*. 2015;26(2):176–192.
6. Hodneland CD, Lee Y-S, Min D-H, Mrksich M. Selective immobilization of proteins to self-assembled monolayers presenting active site-directed capture ligands. *Proc Natl Acad Sci U S A*. 2002;99(8):5048–5052.
7. Galbiati E, Cassani M, Verderio P, et al. Peptide-nanoparticle ligation mediated by cutinase fusion for the development of cancer cell-targeted nanoconjugates. *Bioconjug Chem*. 2015;26(4):680–689.
8. Keppler A, Gendreizig S, Gronemeyer T, Pick H, Vogel H, Johnsson K. A general method for the covalent labeling of fusion proteins with small molecules *in vivo*. *Nat Biotechnol*. 2003;21(1):86–89.
9. Lukinavičius G, Reymond L, Johnsson K. Fluorescent labeling of SNAP-tagged proteins in cells. *Methods Mol Biol*. 2015;1266:107–118.
10. Gautier A, Juillerat A, Heinis C, et al. An engineered protein tag for multiprotein labeling in living cells. *Chem Biol*. 2008;15(2):128–136.
11. Bojkowska K, Santoni De Sio F, Barde I, et al. Measuring *in vivo* protein half-life. *Chem Biol*. 2011;18(6):805–815.
12. Yang G, de Castro Reis F, Sundukova M, et al. Genetic targeting of chemical indicators *in vivo*. *Nat Methods*. 2015;12(2):137–139.
13. Los GV, Encell LP, McDougall MG, et al. HaloTag: a novel protein labeling technology for cell imaging and protein analysis. *ACS Chem Biol*. 2008;3(6):373–382.
14. Raina K, Noblin DJ, Serebrenik YV, Adams A, Zhao C, Crews CM. Targeted protein destabilization reveals an estrogen-mediated ER stress response. *Nat Chem Biol*. 2014;10(11):957–962.
15. Lochte S, Waichman S, Beutel O, You C, Piehler J. Live cell micropatterning reveals the dynamics of signaling complexes at the plasma membrane. *J Cell Biol*. 2014;207(3):407–418.
16. Brigham JL, Perera BGK, Maly DJ. A hexylchloride-based catch-and-release system for chemical proteomic applications. *ACS Chem Biol*. 2013;8(4):691–699.
17. Jing C, Cornish VW. A fluorogenic TMP-tag for high signal-to-background intracellular live cell imaging. *ACS Chem Biol*. 2013;8(8):1704–1712.
18. Szent-Gyorgyi C, Schmidt BF, Creeger Y, et al. Fluorogen-activating single-chain antibodies for imaging cell surface proteins. *Nat Biotechnol*. 2008;26(2):235–240.
19. Telmer CA, Verma R, Teng H, Andreko S, Law L, Bruchez MP. Rapid, specific, no-wash, far-red fluorogen activation in subcellular compartments by targeted fluorogen activating proteins. *ACS Chem Biol*. 2015;10(5):1239–1246.
20. Schwartz SL, Yan Q, Telmer CA, Lidke KA, Bruchez MP, Lidke DS. Fluorogen-activating proteins provide tunable labeling densities for tracking FcεRI independent of IgE. *ACS Chem Biol*. 2015;10(2):539–546.
21. New England Bio markets Snap and Clip tag reagents; Promega markets HaloTag reagents; DHFR-TMP reagents are marketed by Active Motif under LigandLink; FAP and related reagents are marketed by Sharp Edge Labs.
22. Chen I, Howarth M, Lin W, Ting AY. Site-specific labeling of cell surface proteins with biophysical probes using biotin ligase. *Nat Methods*. 2005;2(2):99–104.
23. Eisenstein M. Helping cells to tell a colorful tale. *Nat Methods*. 2006;3(8):647–655.
24. Fernández-Suárez M, Baruah H, Martínez-Hernández L, et al. Redirecting lipoic acid ligase for cell surface protein labeling with small-molecule probes. *Nat Biotechnol*. 2007;25(12):1483–1487.
25. Uttamapinant C, White KA, Baruah H, et al. A fluorophore ligase for site-specific protein labeling inside living cells. *Proc Natl Acad Sci U S A*. 2010;107(24):10914–10919.
26. Uttamapinant C, Sanchez MI, Liu DS, Yao JZ, Ting AY. Site-specific protein labeling using PRIME and chelation-assisted click chemistry. *Nat Protoc*. 2013;8(8):1620–1634.
27. Yin J, Straight PD, McLoughlin SM, et al. Genetically encoded short peptide tag for versatile protein labeling by Sfp phosphopantetheinyl transferase. *Proc Natl Acad Sci U S A*. 2005;102(44):15815–15820.
28. Kawamura T, Stephens B, Qin L, et al. A general method for site specific fluorescent labeling of recombinant chemokines. *PLoS One*. 2014;9(1):1–10.
29. Geierstanger BH, Grunewald J, Bursulaya B. Site-specific labeling methods and molecules produced thereby. WO2013184514A1, 2013.
30. Rashidian M, Song JM, Pricer RE, Distefano MD. Chemoenzymatic reversible immobilization and labeling of proteins without prior purification. *J Am Chem Soc*. 2012;134(20):8455–8467.
31. Mahmoodi MM, Rashidian M, Dozier JK, Distefano MD. Chemoenzymatic site-specific reversible immobilization and labeling of proteins from crude cellular extract without prior purification using oxime and hydrazine ligation. *Curr Protoc Chem Biol*. 2013;5(2):89–109.
32. Dozier JK, Khatwani SL, Wollack JW, Wang YC, Schmidt-Dannert C, Distefano MD. Engineering protein farnesyltransferase for enzymatic protein labeling applications. *Bioconjug Chem*. 2014;25(7):1203–1212.
33. Kim Y, Park T, Woo S, et al. Protein-active agent conjugates and method for preparing the same. WO2012153193A2, 2012.
34. Heal WP, Wickramasinghe SR, Bowyer PW, et al. Site-specific N-terminal labelling of proteins *in vitro* and *in vivo* using N-myristoyl transferase and bioorthogonal ligation chemistry. *Chem Commun (Camb)*. 2008;3(4):480–482.
35. Heal WP, Wickramasinghe SR, Leatherbarrow RJ, Tate EW. N-Myristoyltransferase-mediated protein labelling in vivo. *Org Biomol Chem*. 2008;6(13):2308–2315.
36. Kulkarni C, Kinzer-Ursem TL, Tirrell DA. Selective functionalization of the protein N terminus with N-myristoyl transferase for bioconjugation in cell lysate. *Chembiochem*. 2013;14(15):1958–1962.
37. Dennler P, Chiotellis A, Fischer E, et al. Transglutaminase-based chemoenzymatic conjugation approach yields homogeneous antibody-drug conjugates. *Bioconjug Chem*. 2014;25(3):569–578.
38. Dennler P, Schibli R, Fischer E. Enzymatic antibody modification by bacterial transglutaminase. *Methods Mol Biol*. 2013;1045:205–215.
39. Lhospice F, Brégeon D, Belmant C, et al. Site-specific conjugation of monomethyl Auristatin E to Anti-CD30 antibodies improves their pharmacokinetics and therapeutic index in rodent models. *Mol Pharm*. 2015;12(6):1863–1871.
40. Strop P, Liu SH, Dorywalska M, et al. Location matters: site of conjugation modulates stability and pharmacokinetics of antibody drug conjugates. *Chem Biol*. 2013;20(2):161–167.
41. Dorywalska M, Strop P, Melton-Witt JA, et al. Effect of attachment site on stability of cleavable antibody drug conjugates. *Bioconjug Chem*. 2015;26(4):650–659.
42. Oral presentation by Hans-Peter Gerber, Porto, Portugal, April 13, 2015.
43. Popp MW, Antos JM, Grotenbreg GM, Spooner E, Ploegh HL. Sortagging: a versatile method for protein labeling. *Nat Chem Biol*. 2007;3(11):707–708.
44. Guimaraes CP, Witte MD, Theile CS, et al. Site-specific C-terminal and internal loop labeling of proteins using sortase-mediated reactions. *Nat Protoc*. 2013;8(9):1787–1799.
45. Schmohl L, Schwarzer D. Sortase-mediated ligations for the site-specific modification of proteins. *Curr Opin Chem Biol*. 2014;22:122–128.

46. Ling JJ, Policarpo RL, Rabideau AE, Liao X, Pentelute BL. Protein thioester synthesis enabled by sortase. *J Am Chem Soc*. 2012;134(26):10749–10752.
47. Zhang C, Dai P, Spokoyny AM, Pentelute BL. Enzyme-catalyzed macrocyclization of long unprotected peptides. *Org Lett*. 2014;16(14):3652–3655.
48. Paterson BM, Alt K, Jeffery CM, et al. Enzyme-mediated site-specific bioconjugation of metal complexes to proteins: sortase-mediated coupling of copper-64 to a single-chain antibody. *Angew Chem Int Ed Engl*. 2014;53(24):6115–6119.
49. Policarpo RL, Kang H, Liao X, Rabideau AE, Simon MD, Pentelute BL. Flow-based enzymatic ligation by sortase A. *Angew Chem Int Ed Engl*. 2014;53: 9203–9208.
50. Chen I, Dorr BM, Liu DR. A general strategy for the evolution of bond-forming enzymes using yeast display. *Proc Natl Acad Sci U S A*. 2011;108(28):11399–11404.
51. Dorr BM, Ham HO, An C, Chaikof EL, Liu DR. Reprogramming the specificity of sortase enzymes. *Proc Natl Acad Sci U S A*. 2014;111(37):13343–13348.
52. Williamson DJ, Fascione MA, Webb ME, Turnbull WB. Efficient N-terminal labeling of proteins by use of sortase. *Angew Chem Int Ed Engl*. 2012;51(37): 9377–9380.
53. Williamson DJ, Webb ME, Turnbull WB. Depsipeptide substrates for sortase-mediated N-terminal protein ligation. *Nat Protoc*. 2014;9(2):253–262.
54. Bellucci JJ, Bhattacharyya J, Chilkoti A. A noncanonical function of sortase enables site-specific conjugation of small molecules to lysine residues in proteins. *Angew Chem Int Ed Engl*. 2014;54:441–445.
55. Jain N, Smith SW, Ghone S, Tomczuk B. Current ADC linker chemistry. *Pharm Res*. March 11, 2015.
56. Ageitos JM, Baker PJ, Sugahara M, Numata K. Proteinase K-catalyzed synthesis of linear and star oligo(L-phenylalanine) conjugates. *Biomacromolecules*. 2013; 14(10):3635–3642.
57. Numata K, Baker PJ. Synthesis of adhesive peptides similar to those found in blue mussel (Mytilus edulis) using papain and tyrosinase. *Biomacromolecules*. 2014;15(8):3206–3212.
58. Liebscher S, Schöpfel M, Aumüller T, et al. N-terminal protein modification by substrate-activated reverse proteolysis. *Angew Chem Int Ed Engl*. 2014;53(11): 3024–3028.
59. Liebscher S, Kornberger P, Fink G, et al. Derivatization of antibody Fab fragments: a designer enzyme for native protein modification. *Chembiochem*. 2014; 15(8):1096–1100.
60. Poster from PEGS Boston 2014; WO 2006/015879 A1. Available at: www.eucodis.com
61. Zakeri B, Fierer JO, Celik E, et al. PNAS Plus: peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proc Natl Acad Sci U S A*. 2012;109(12):E690–E697.
62. Sun F, Zhang W, Mahdavi A, Arnold FH, Tirrell DA. Synthesis of bioactive protein hydrogels by genetically encoded SpyTag-SpyCatcher chemistry. *Proc Natl Acad Sci U S A*. 2014;2014(31):1–6.
63. Botyanszki Z, Tay PKR, Nguyen PQ, Nussbaumer MG, Joshi NS. Engineered catalytic biofilms: site-specific enzyme immobilization onto *E. coli* curli nanofibers. *Biotechnol Bioeng*. 2015;9999:1–9.
64. Fierer JO, Veggiani G, Howarth M. SpyLigase peptide-peptide ligation polymerizes affibodies to enhance magnetic cancer cell capture. *Proc Natl Acad Sci U S A*. 2014;111(13):E1176–E1181.
65. Appel MJ, Bertozzi CR. Formylglycine, a post-translationally generated residue with unique catalytic capabilities and biotechnology applications. *ACS Chem Biol*. 2015;10(1):72–84.
66. Smith EL, Giddens JP, Iavarone AT, Godula K, Wang LX, Bertozzi CR. Chemoenzymatic Fc glycosylation via engineered aldehyde tags. *Bioconjug Chem*. 2014;25(4):788–795.
67. Agarwal P, Kudirka R, Albers AE, et al. Hydrazino-pictet-spengler ligation as a biocompatible method for the generation of stable protein conjugates. *Bioconjug Chem*. 2013;24(6):846–851.
68. Kudirka R, Barfield RM, McFarland J, et al. Generating site-specifically modified proteins via a versatile and stable nucleophilic carbon ligation. *Chem Biol*. 2015;22(2):293–298.
69. Holder PG, Jones LC, Drake PM, et al. Reconstitution of formylglycine-generating enzyme with copper(II) for aldehyde tag conversion. *J Biol Chem*. 2015;290(25):15730–15745.
70. Long MJC, Hedstrom L. Mushroom tyrosinase oxidizes tyrosine-rich sequences to allow selective protein functionalization. *Chembiochem*. 2012;13(12):1818–1825.
71. Wals K, Ovaa H. Unnatural amino acid incorporation in *E. coli*: current and future applications in the design of therapeutic proteins. *Front Chem*. 2014;2:15.
72. Tian F, Lu Y, Manibusan A, et al. A general approach to site-specific antibody drug conjugates. *Proc Natl Acad Sci U S A*. 2014;111(5):1766–1771.
73. Kularatne SA, Deshmukh V, Ma J, et al. A CXCR4-targeted site-specific antibody-drug conjugate. *Angew Chem Int Ed Engl*. 2014;53(44):11863–11867.
74. Zimmerman ES, Heibeck TH, Gill A, et al. Production of site-specific antibody-drug conjugates using optimized non-natural amino acids in a cell-free expression system. *Bioconjug Chem*. 2014;25(2):351–361.
75. Palcic MM. Glycosyltransferases as biocatalysts. *Curr Opin Chem Biol*. 2011; 15(2):226–233.
76. Hudak JE, Bertozzi CR. Glycotherapy: new advances inspire a reemergence of glycans in medicine. *Chem Biol*. 2014;21(1):16–37.
77. Villalonga ML, Díez P, Sánchez A, Gamella M, Pingarrón JM, Villalonga R. Neoglycoenzymes. *Chem Rev*. 2014;114(9):4868–4917.
78. Wang LX, Amin MN. Chemical and chemoenzymatic synthesis of glycoproteins for deciphering functions. *Chem Biol*. 2014;21(1):51–66.
79. Okamoto R, Izumi M, Kajihara Y. Decoration of proteins with sugar chains: recent advances in glycoprotein synthesis. *Curr Opin Chem Biol*. 2014;22:92–99.
80. Rouhanifard SH, López-Aguilar A, Wu P. CHoMP: a chemoenzymatic histology method using clickable probes. *Chembiochem*. 2014;15(18):2667–2673.
81. Zeglis BM, Davis CB, Abdel-Atti D, et al. Chemoenzymatic strategy for the synthesis of site-specifically labeled immunoconjugates for multimodal PET and optical imaging. *Bioconjug Chem*. 2014;25(12):2123–2128.
82. Zou G, Ochiai H, Huang W, Yang Q, Li C, Wang LX. Chemoenzymatic synthesis and Fcγ receptor binding of homogeneous glycoforms of antibody Fc domain. Presence of a bisecting sugar moiety enhances the affinity of Fc to FcγIIIa receptor. *J Am Chem Soc*. 2011;133(46):18975–18991.
83. Huang W, Giddens J, Fan SQ, Toonstra C, Wang LX. Chemoenzymatic glycoengineering of intact IgG antibodies for gain of functions. *J Am Chem Soc*. 2012; 134(29):12308–12318.
84. Schwarz F, Huang W, Li C, et al. A combined method for producing homogeneous glycoproteins with eukaryotic N-glycosylation. *Nat Chem Biol*. 2010;6(4): 264–266.
85. Valderrama-Rincon JD, Fisher AC, Merritt JH, et al. An engineered eukaryotic protein glycosylation pathway in *Escherichia coli*. *Nat Chem Biol*. 2012;8(5): 434–436.
86. Merritt JH, Fisher AC, Hamilton BS, DeLisa MP. Polysialic acid, blood group antigens and glycoprotein expression in prokaryotes. US20140273103, 2014.
87. Van Delft FL, Van Geel R, Wijdeven MA. Modified antibody, antibody-Conjugate and process for the preparation thereof. WO2014065661A1, 2014.
88. Satomaa T, Helin J, Ekholm FS. Conjugates of a glycoprotein or a glycan with a toxic payload. WO2014177771A1, 2014.
89. Klausen NK, Bjorn S, Behrens C, Garibay, PW. Pegylated Factor VII Glycoforms. US20090227504, 2009.
90. Tiels P, Baranova E, Piens K, et al. A bacterial glycosidase enables mannose-6-phosphate modification and improved cellular uptake of yeast-produced recombinant human lysosomal enzymes. *Nat Biotechnol*. 2012;30(12):1225–1231.